(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 December 2001 (20.12.2001)

PCT

(10) International Publication Number
WO 01/97134 A1

(51) International Patent Classification[7]: G06F 17/60

(21) International Application Number: PCT/US01/18076

(22) International Filing Date: 5 June 2001 (05.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/209,936 7 June 2000 (07.06.2000) US

(71) Applicant and
(72) Inventor: **BARBER, Timothy, P.** [US/US]; 120 North Mobley Drive, Boise, ID 83712 (US).

(74) Agent: **BROWN, Mark, E.**; Shughart Thomson & Kilroy, P.C., Suite 1800, Twelve Wyandotte Plaza, 120 West 12th Street, Kansas City, MO 64105 (US).

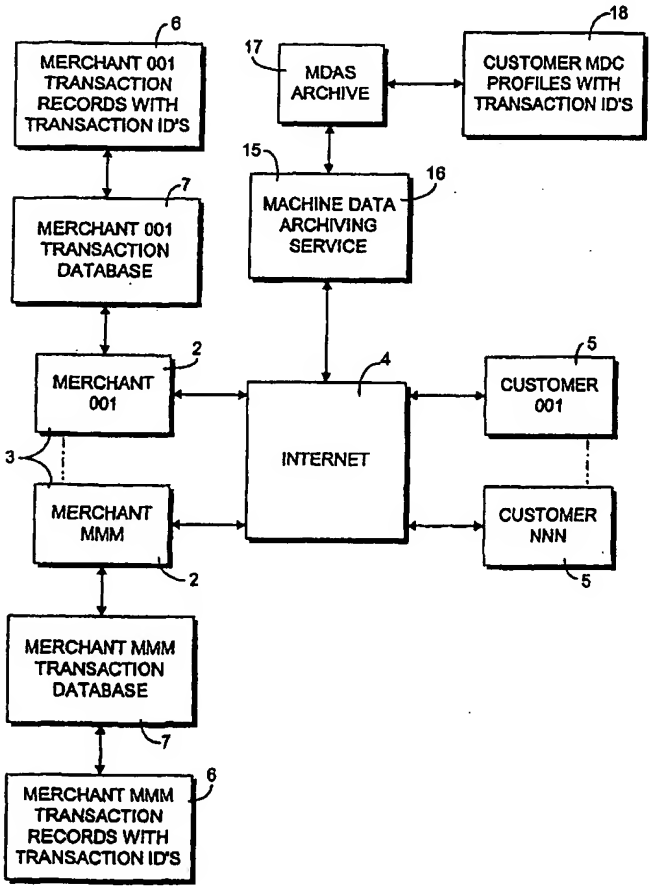
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

[Continued on next page]

(54) Title: ONLINE MACHINE DATA COLLECTION AND ARCHIVING PROCESS



(57) Abstract: An online machine data collection and archiving process (15) generates a machine data profile (18) of a customer computer (5) accessing a transaction form of a merchant web site (3) and links the machine data profile (18) and a transaction record (6) with customer identifying information using a unique transaction identification string. The process preferably captures parameters typically communicated as part of web accesses, such as an IP address, an HTTP header, and cookie information. The process additionally causes the customer computer (5) to process self-identification routines by processing coding within the merchant transaction form, the self-identification routines yielding further profile parameters. The process further includes a routine for bypassing an intervening proxy to the merchant web site (3) to reveal the true IP address of the customer computer (5).

WO 01/97134 A1

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

ONLINE MACHINE DATA COLLECTION AND ARCHIVING PROCESS

Cross-Reference to Provisional Application

This application claims benefits from provisional application, Serial No. 60/209,936 filed June 7, 2000 and entitled METHOD FOR COLLECTING MACHINE DATA FROM CUSTOMERS ON A COMPUTER NETWORK.

Background of the Invention

The present invention relates to identity detection techniques and, more particularly, to a process for collecting and utilizing machine-identifying data of computers and other online appliances used in online interactions and transactions and associating the collected machine data with such online interactions.

The internet, or global computer network, represents a new medium for marketing similar to the way mail ordering and telephone ordering did in the past. A downside of internet marketing is that it also presents new opportunities for unscrupulous persons to take advantage of the mechanisms of internet transactions by fraudulent and deceptive practices. Merchants and financial institutions bear the initial costs of fraud. However, consumers ultimately pay the costs in the form of prices and credit rates which must take into account losses from fraud. Internet purchases typically involve the use of web page forms which are filled in by the customer with identity, address, purchase, shipping, and

payment information and submitted to the online merchant for processing. Internet purchases are most often paid for by way of credit cards. While a merchant's software may be able to verify the existence and status of a credit card account number and an authorization for a specific amount, the merchant is often not able to match a credit card number with a specific purchaser or shipping address. Thus, absent any overt indication otherwise, a merchant generally assumes that anyone using a credit card is authorized to do so and that a customer is who he identifies himself to be.

An important step in combating fraud is accurate identification of the computers through which customers make transactions and associating such identities with transactions which arouse suspicions or which ultimately turn out to be fraudulent. Basic machine identity is essential to the manner in which the internet operates. We speak in terms of "going" to a web site. In reality, "going" to a web site involves sending a request for a web page file in a directory or folder on a computer located at a specific internet protocol, or IP, address. In order for the web page file to be returned to the requesting computer for processing into a displayed "web page", the request must include return "directions" in the form of the basic identity of the requesting computer, including its IP address. Some web sites are implemented with software which enables responses to web page requests to be tailored to specifics of the requesting machine's configuration, specific web browser, and the like. For this reason, current versions of browsers usually communicate configuration information in addition to a return IP address and return path.

The IP address of a page requesting computer can give an indication of the specific country where the computer is located. Further, identification of a page-requesting computer can also recognize a returning user using the same computer as during a previous access. For example, placing an HTTP (hypertext transfer protocol) "cookie" on a page-requesting computer can make it possible to identify the computer on a later access.

Because direct interaction with a customer's computer is essential in detecting fraud, it has been assumed that any viable fraud detection software must be integrated with a merchant's software. As a result, most existing fraud detection solutions require merchants to either abandon or extensively modify their existing web-based transaction processing software. An additional problem with focusing fraud detection at single merchants is that perpetrators of fraud often hit many merchants in an attempt to avoid or delay detection. Thus, an ideal system for fraud detection in online marketing would only minimally affect the merchant's existing software and would route fraud detection efforts through a central, third-party entity serving a large multitude of merchants.

Summary of the Invention

The present invention provides a process for collecting data associated with a customer's computer during access of a merchant, financial, other host web site, and associating a transaction identification number with the data and with a transaction form of the merchant. Generally, the present invention captures machine identifying data from a

computer or other digital appliance accessing a host web site, sends the captured data to a machine data archive along with a unique transaction identification string for storage in the archive and writes the same transaction identification string into a transaction form through which transactions with the host web site are conducted. The machine data is, thus, associated with the customer identification data within the transaction form by way of the transaction identification string and can be used on-the-fly or at a later time for a variety of purposes including, but not limited to, fraud detection. Although the term "archive" is used, the machine data collected need not be stored permanently.

The machine data collection process of the present invention is intended to be employed in a variety of applications including, but not limited to: online purchases and orders; online banking, bill payment, and funds transfers; online registrations, such as for memberships, product warranties, applications for credit, renewal of subscriptions and licenses; online technical support; and the like. The term "transaction" is used in the present invention to describe an interaction effected between a digital appliance and a host system. However, the term "transaction" is not intended to be restricted to only commercial interactions involving purchases. The term "transaction" is intended to apply to an interaction of a remote digital appliance with a host system using a relatively anonymous type of access process over a digital medium in which some form of self-identification of the accessing appliance is inherent in the access process and in which the true identity of the accessing party, the true source address of the appliance on the

medium, and/or the true machine characteristics of the accessing appliance is/are essential or desirable to the interaction.

The host entity which operates the host system accessed is intended to encompass a commercial, financial, educational, governmental, associational, or other type of entity. The term "merchant" will be used herein to refer to such a host entity without intent to limit the present invention to commercial transactions. The medium of access is intended to be interpreted as including a global computer network such as the internet or world wide web, as well as other types of networks which may be less than global but which are publicly and/or anonymously accessible. The term "internet" will be used herein to refer to the medium through which accesses to the host entity are made. The terms "customer computer" or "machine" are used herein to refer to a device for effecting remote access to a host system over a digital medium and are meant to encompass not only conventional types of personal computers, but also additional types of "digital appliances" with online access capabilities, such as: cell phones, personal digital assistant devices, electronic game systems, television sets with online access capabilities, web appliances for vehicles, and any other type of device with online access capabilities whether connected to a wired communications network directly or by a radiant technology.

The machine data collection process of the present invention contemplates a two party process embodiment in which a "merchant" processes and/or stores machine data profiles of customer computers in-house, as well as a three party process embodiment in

which machine data profiles of customer computers are processed and/or stored for the merchant by a third-party machine data collection and archive service.

In a two party embodiment of the data collection process of the present invention, the customer machine data is captured by a merchant or host system which also generates a unique transaction identification (ID) string and assigns or associates the transaction ID with a machine data profile of the customer machine data profile. In the two party process, the merchant system captures customer computer data which is inherently passed from the customer computer to the merchant's web site, such as an IP address of the customer computer and an HTTP header. Additionally, according to the present invention, the merchant web page code may have routines or calls for external routines which, when processed by the customer computer, cause the customer computer to further identify itself by collecting and returning certain machine and software configuration characteristics, which can be used to identify the particular customer computer. The two party process may include the generation and setting of an HTTP cookie in the customer browser for recognition upon a later access with the merchant web site.

Although the two-party embodiment of the machine data collection process of the present invention has utility for some applications, the three party embodiment is preferred for applications in which analysis of a maximum number of customer computer profiles is desirable, such as certain types of marketing processes and fraud detection and control processes. In a three embodiment of the present invention, the customer machine data of

computers accessing the second party or merchant web site is communicated to and stored within a third party system, referred to herein as a machine data archive service. In the three party process, the transaction ID could be generated by the merchant system, but is preferably generated by the archive service. The use of the term "archive" is not meant to indicate that the customer machine data profiles are stored permanently within the third party system. Permanent storage of such profiles may not be practical, as far as yielding beneficial results to the purposes for which the profiles are collected. Thus, the term "archive" is meant to indicate a central storage facility, such as a database, with a selected retention period, with purging of most profiles after a certain length of time.

In the three party process a routine or line of code is added into the hypertext markup language (HTML) code which defines the merchant's web page, particularly an order or transaction form page. The added routine issues a request for a machine data collection (MDC) script to the third-party web site when the form page code is processed by the customer's browser. When the script request is received by the machine data archiving service (MDAS), the archive service generates a unique transaction identification (TA/ID) and checks for its own cookie. If no MDAS cookie is present, the archive service sends a cookie to the requesting computer along with a machine data collection (MDC) script having the transaction ID embedded therein. The MDC script is executed by the customer's browser, causing collection of certain data from the customer's computer which is sent back to the archive service along with the transaction ID and stored in a machine data profile in the machine data archive. The transaction ID is

written into the transaction form, and when the transaction form is submitted to the merchant web site, the transaction ID string becomes a part of the transaction data record, along with customer identification, location, and financial information.

The machine data initially collected and stored in each profile preferably includes the transaction ID, the apparent IP address of the customer's computer, a conventional HTTP header which identifies the customer's browser versions and certain configuration aspects of the browser, and the archive service's cookie. The combination of such information, minus the transaction ID, will be relatively rare but may not be unique. Additionally, customers intent on conducting fraudulent transactions often hide their IP address behind HTTP proxies. In order to further narrow the machine profile, in a preferred embodiment of the present invention, the MDC script performs additional machine profiling operations: generation of a machine "fingerprint" and a proxy "piercing" operation.

In the fingerprint generation operation, the MDC script assembles an attribute string formed by various attributes or configuration settings of the browser which can be queried by the script. The MDC script then performs a conversion process on the attribute string to generate a fingerprint string having content which is a function of the original content of the attribute string. The conversion process is preferably a "hashing" function which is, in effect, an irreversible encryption algorithm. The generation of a conventional checksum is one example of a type of hashing function. For example, if the attribute string is formed by alphanumeric characters, the conversion process is performed

on the string of codes representing the characters. The particular conversion process or hashing function used may be one of many types of conventional conversion algorithms or hashing functions, which are typically used for data integrity tests. The resulting string from the MDC conversion process is a so-called fingerprint, which is returned to the archive service along with the transaction ID for storage in the machine profile. A time value, queried from the customer computer time-of-day clock, is returned with the fingerprint string and stored in conjunction therewith.

An HTTP proxy is one of several types of proxies through which a browser may be setup to operate. Setting up an HTTP proxy causes HTTP requests to be relayed by a primary gateway, through which the computer actually interfaces to the internet, to a remote secondary gateway, or proxy, with an IP address different from the primary gateway IP address. Such redirection hides the true IP address of a computer. The proxy piercing operation of the MDC script queries the customer computer for any LAN (local area network) address which may be assigned to the computer and reads the system time of day clock. Then attempts are made to send the LAN address, if any, the time value, and the transaction ID to the archive service using a protocol which will not be redirected through the HTTP proxy, for example a lower level protocol such as TCP/IP or UDP protocols. If the attempt is successful, the message containing the time value, the transaction ID, and LAN address arrive at the archive service web site with the true return IP address of the customer computer, whether an HTTP proxy intervenes or not. The LAN address and IP address so derived are stored in the machine profile. It should be

noted that the use of an HTTP proxy is not, of itself, an indication of fraud. However, the acquisition of an additional IP address is one more parameter with which to identify a particular computer.

When, and if, the customer submits the transaction form to the merchant, the transaction ID string is communicated to the merchant, along with other customer information such as name, address, credit card number and the like plus transaction information. The complete transaction record is stored on the merchant's system and is associated with a specific machine identity profile within the archive service by way of the transaction ID string. Thereafter, the stored machine identity profiles and transaction records of large numbers of transactions can be analyzed by various fraud detection techniques to detect patterns of fraud and fraud attempts and, preferably, identify and locate the sources of such activity.

The machine data profiles stored in the archive service need not be combined with the customer identification information for non-suspicious transactions, to thereby preserve the privacy of non-suspicious customers within the machine data archive. However, the processes of the present invention do not require that the customer identification information be kept separated from any associated machine data profiles, and there may be reasons to combine the associated records.

Brief Description of the Drawings

Fig. 1 is a simplified block diagram illustrating a plurality of customer and merchant computers interfaced to the internet along with a machine data archiving service computer for practicing the machine data collection process of the present invention.

Fig. 2 is a simplified block diagram illustrating connection of a customer computer to the internet, with optional components shown in phantom lines.

Fig. 3 is a flow diagram illustrating principal steps of the machine data collection and archiving process according to the present invention.

Fig. 4 is a flow diagram illustrating more detailed steps of the machine data collection and archiving process according to the present invention.

Fig. 5 is a flow diagram illustrating a still further detailed steps in the machine data collection and archiving process of the present invention.

Various objects and advantages of this invention will become apparent from the following description taken in relation to the accompanying drawings wherein are set forth, by way of illustration and example, certain embodiments of this invention.

The drawings constitute a part of this specification, include exemplary embodiments of the present invention, and illustrate various objects and features thereof.

Detailed Description of the Invention

As required, detailed embodiments of the present invention are disclosed herein; however, it is to be understood that the disclosed embodiments are merely exemplary of the invention, which may be embodied in various forms. Therefore, specific structural and functional details disclosed herein are not to be interpreted as limiting, but merely as a basis for the claims and as a representative basis for teaching one skilled in the art to variously employ the present invention in virtually any appropriately detailed structure.

Referring to the drawings in more detail:

The reference numeral 1 (Fig. 3) generally designates a process for online collection of machine identifying or profiling data of computers involved in commercial transactions and for archiving such data to facilitate analysis for fraud detection purposes. The process collects machine identifying or profiling data of computers involved in commercial transactions and archives such data in a third-party machine data archive service in association with a transaction identification string or ID which is also written into a transaction form of a merchant with whom the customer is conducting a transaction.

Fig. 1 illustrates a plurality of host entities or merchants with corresponding merchant computers 2, on which are operated merchant web sites 3 which are accessible over a global computer network, such as the internet 4, by a plurality of customer computers 5. The merchant computers 2 execute various programs which enable the sale of products or services by way of the internet 4. The merchant web sites 3 typically make

use of form type web pages with which the customers 5 interact by filling in various data fields, for example, name, address, shipping address, telephone number, credit card type and number and expiration date, and description and quantities of products to be ordered.

The merchant transaction forms are usually written in hypertext markup language (HTML) and may include requests for code written in other languages, such as Java and the like. When a customer 5 accesses a merchant's transaction form, a transaction form file is communicated to the customer's computer with various data fields displayed as fill-in boxes or the like. The customer fills in the appropriate fields and selects a submit "button" which activates a routine to transfer the collected information back to the merchant web site 3 for processing. The returned "form" is a data record 6 which is stored in a merchant transaction database 7 for retrieval and processing in due course to cause the ordered items to be gathered, packaged and prepared for shipment, along with financial processing to debit the customer's credit account. The financial processing may include a validity check of the credit account and an authorization check for the amount of purchase with the credit card issuer. Additionally, inventory management processes are executed based on the items withdrawn from stock for shipment.

In a three party embodiment of the present invention, the process 1 makes use of an entity referred to herein as a machine data archiving service, MDAS or archive service, which operates an archive service computer system 15, including an archive service web site 16. The archive service system 15 maintains a machine data archive service database or archive 17 in which the machine data collection profiles 18 from customer computers 5

of the merchants 2 are stored. The archive service web site 16 is interfaced to the internet 4. The archive service 15 is preferably independent of the merchants and may be operated by a merchants' association, a financial institution or association thereof, or may be an independent contractor. Alternatively, it is conceivable that a merchant with a high volume of online sales could operate its own in-house machine data profile collection and archiving service 15, for fraud detection or possibly for marketing purposes.

Referring to Fig. 2, a customer computer system 5 includes a customer computer 20 interfaced to the internet 4 by way of a primary gateway 22, as of an internet service provider (ISP). The computer 20 might be one of many on a local area network or LAN 24 which includes a router or switch which routes data from the internet 4 to the computers on the network. The computer 20 may communicate through the internet 4 by way of a HTTP (hypertext transfer protocol) proxy 26, which disguises the internet protocol (IP) address of the actual gateway 22. The computer 20 accesses web sites on the internet 4 using a customer web browser 28 which processes HTML language and various other standard web oriented languages to display or otherwise render the content of web pages and interact therewith. The browser 28 is normally enabled to accept "cookies" 30 which are stored in a cookie file. Cookies 30 are data strings which are issued by web sites and give an indication of a previous visit to a particular web site and may indicate a particular configuration or set of preferences of the customer's setup of the computer 20. Typically, the customer computer 20 has a time of day clock/calendar 32.

The customer computer 20 may have a fixed IP address, depending on the manner in which it is interfaced to the internet. More commonly, the customer computer 20 will have a temporary or dynamically assigned IP address which is determined by the primary router 22. The primary router 22 has an IP address, as do a router of a LAN 24 or an HTTP proxy 26 if either is present in the customer's computer system 5.

Fig. 3 illustrates the principal actions or steps of a general or basic process 34 of the process 1 for collecting machine identifying data from customer computers 5. At step 35, at least one machine identifying profile parameter is captured upon access of a customer computer 5 or other online access device with a host or merchant web site 3. A unique transaction identifier or TA/ID is generated at 36 and associated at 37 with the captured profile parameter. The transaction ID is also associated at step 38 with a transaction record generated as a result of the interaction or transaction conducted between the customer computer 5 and a merchant web site 3. Although not specifically shown in Fig. 3, the process 34 may capture machine profile data that is passed from the customer computer 5 to the merchant computer 3 as an inherent step of the customer computer 5 accessing the merchant computer 3. Alternatively, the process 34 may pass routines to the customer computer 5 to cause it to "self-identify" itself by querying certain configuration parameters and passing such information to a machine profile stored either within the merchant's system 2 or in a third party archive 17. The process 34, thus, encompasses a two-party embodiment or a three party embodiment of the machine data collection and archiving process 1 of the present invention.

Referring particularly to Fig. 4, a more particular three party embodiment of the machine data collection and archiving process 1 begins at step 40 with the coding of a machine data collection (MDC) script request into the web page code for a transaction form of a merchant web site 3. When a customer 5 accesses the merchant transaction form at step 42, the customer browser 28 processes the transaction page code, including the MDC script request, which causes the MDC script request to be communicated to the archive service web site 16 at step 44. The script request arrives at the archive service 15 with a set of customer machine parameters which principally provide a return path for the MDC script from the archive service 15 to the customer 5. The customer machine parameter set preferably includes "user agent" information, which is the version of the customer browser 28.

At step 46, the archive service 15 generates a unique transaction ID string and associates it with the customer machine parameter set in the MDAS archive 17. At step 48, the archive service returns the MDC script, with the transaction ID embedded within it, to the customer browser 28. At step 50, the customer browser 28 processes the MDC script which, at a minimum, writes the transaction ID string into the merchant's transaction form. Assuming that the customer 5 completes the transaction and submits the transaction form to the merchant 2 at step 52, the transaction ID string is stored with the transaction data record 6 in the merchant transaction database 7. The transaction ID, thus, indirectly associates the machine data parameter set 18 stored in the MDAS archive 17 at step 54 with the customer identity information stored with the transaction data record 6 in the

merchant's transaction database 7. Thereafter, qualified parties may access the MDAS archive 17 for information related to a transaction ID.

The MDAS archive 17 need not contain any information which specifically identifies a particular customer, only the machine parameter profiles 18 with associated transaction ID strings. The MDAS archive records 18 can be analyzed in conjunction with the merchant transaction records for patterns of fraud or for other purposes. The great majority of MDAS archive records can be purged from the archive 17 after a selected period of time. Any records which are associated with any transaction irregularities or suspicions of actual fraud may be retained longer.

Fig. 5 illustrates the principal steps of a preferred embodiment 60 of the machine data collection and archiving process 1 of the present invention. The process 60 begins with the addition at 62 of a machine data collection (MDC) script to the transaction (TA) form page code of a merchant web site 3. The transaction form page code is processed at 64 by a customer browser 28 when the merchant web page is accessed to thereby request the MDC script at 66 from the Machine data archive service (MDAS) web site 16. When the browser 28 accesses the MDAS web site 16, requesting the MDC script, the MDAS web site checks for the presence of an MDAS cookie at step 68. If no MDAS cookie is detected, an MDAS cookie is generated at 70 and a unique transaction identification (TA/ID) string is generated at 72. The MDC script, transaction ID, and cookie, if not previously set, are returned at 74 to the customer browser 28, the transaction ID being embedded within the MDC script.

When the MDC script is received by the browser 28, it is executed at 76. The cookie is stored in the cookie file 30, or possibly in the memory of the customer computer 20. Execution of a preferred MDC script causes several actions to be performed. The MDC script writes the transaction ID into the transaction form at step 78. The script can do this by either setting an existing variable of an appropriate name to the transaction ID string or by writing an appropriate variable into the transaction form page and setting its value to the transaction ID string. Additionally, the preferred MDC script generates a "fingerprint" of the customer computer 20 at step 80 and attempts to perform a proxy piercing operation at step 82.

In generating the machine fingerprint at 80, the MDC script queries the browser 28 for a number of attributes and settings and concatenates the results into an attribute string at 84. The MDC script then performs a hashing algorithm on the attribute string at 80 to generate a fingerprint string which has a high degree of uniqueness. Hashing functions are irreversible encryption processes in which the result is dependent on the original content of the data on which the hashing algorithm is operated. Hashing functions are commonly used for data integrity checking. As previously stated, a common checksum is the result of a type of hashing function. The particular hashing function employed preferably maximizes the uniqueness of the resulting fingerprint.

At step 86, the customer computer clock 32 is queried for a current time value. At step 88, the fingerprint, the transaction ID, and the time value are communicated to the

MDAS web site 16 along with an HTTP header with cookie and "apparent" IP address, all of which are stored as a machine data profile 18 within the MDAS archive 17.

At step 90, the MDC script adds a proxy piercer request to the transaction form which, when executed by the browser 28 at step 92, sends a request for a proxy piercer applet or code to the MDAS web site 16. When the proxy piercer applet/code is executed by the browser 28 at 94, a time value from the clock 32 is again queried at 96 and any existing local area network (LAN) address is queried at 98. At step 100, the proxy piercer applet/code sends the time value, the LAN address (if any), and the transaction ID to the MDAS web site 16 by a protocol which bypasses any existing HTTP proxy 26. The protocol used is one which is at a lower level than HTTP, such as UDP (user datagram protocol) or, preferably, TCP/IP (transmission control protocol/internet protocol).

Bypassing the HTTP proxy 26 causes the data sent in step 100 to arrive at the MDAS web site 16 with the IP address of the primary gateway 22, which may be different from any apparent IP address previously recorded if an HTTP proxy 26 intervenes. If the proxy piercer procedure 82 is successful, the primary gateway IP address is stored at step 102 within the machine data profile 18 identified by the transaction ID. It should be noted that some types of proxies, such as some types of firewalls, may block all non-HTTP protocol packets, so that the proxy piercer procedure 82 might not be successful in all cases.

If the customer completes the transaction with the merchant web site 3, the transaction form is submitted at step 104, which causes the transaction record 6, including the transaction ID, to be stored at step 106 in the merchant database 7 for processing.

Following are examples of code for an MDC script, as from steps 40 or 62. Assuming the machine data archiving service or MDAS web site 16 has the fictional URL (uniform resource locator) example-url.net and a specific merchant has a merchant identifier MMM, a line of HTML code is added at step 40 to the transaction form of merchant MMM between the <form> and </form> HTML tags which has the form:

<script src=https://www.example-url.net/s/?MMM></script>

When the customer browser 28 processes the transaction form at step 42, it requests a script file from the source URL: https://www.example-url.net/s/?MMM.

At step 44, the customer web browser 28 requests the MDC script by way of the HTTP protocol. The HTTP request includes the merchant ID MMM, the user agent (browser version), the IP address of the customer's HTTP proxy, and any HTTP cookies previously sent to the customer by www.example-url.net. Upon receiving this information, the archive service 16 records this information in a machine data record which also includes the transaction ID.

Upon receiving the file request, the archive service 16 generates a unique transaction ID (represented below as ZZZ) at step 46 to be associated with the transaction

and the machine parameter set. An exemplary transaction ID is a string of 24 letters and digits. The first eight digits form a time-stamp which is a hexadecimal representation of the seconds elapsed since midnight January 1, 1970 UTC (coordinated universal time).

In the preferred embodiment of the process 1, the MDC script is written in an ECMAScript compliant language, such as JavaScript, JScript, or VBScript. A JavaScript version of the MDC script is as follows (linebreaks and indentations added for clarity):

```
document.write("<input name=transactionid type=hidden value=ZZZ>;

d=new Date();

t=3600*d.getHours()+60*d.getMinutes()+d.getSeconds();

document.write("<img height=1 width=1      src=https://www.example-
        url.net/t/?i=ZZZ&t="+t+">");

document.write("<applet height=1 width=1
codebase=https://www.example-url.net/
code=a/?ZZZ>
<param name=i value=ZZZ></applet>");
```

The exemplary MDC script includes the unique transaction ID value in several places. When the script executes on the customer computer 20, it writes HTML code into the merchant's order form. Specifically:

1) The script adds a hidden variable called "transactionid" to the merchant's transaction form and assigns it the value of the transaction ID (ZZZ). When the transaction form is submitted, the merchant receives the transaction ID and can associate it with the transaction data record.

2) The script computes the seconds elapsed since midnight on the clock 32 and writes a request for a 1 pixel by 1 pixel image. Included in the request is the transaction ID and the time value. When the request executes, this data is sent back to the archive service 16 and recorded with the transaction ID in the MDAS archive 17.

3) The script adds a request for a program located at the archive service web site 16 which, in this example, is a Java applet. The applet downloads to the customer computer 20 from the archive service 16 and executes, appearing as a 1 pixel by 1 pixel image on the transaction form. The transaction ID is passed to the program as a parameter specified in the script. The program performs three tasks:

a) it calculates TTT, the seconds elapsed since midnight on the system clock 32;

b) it calculates AAA, the address of the customer 20 on its own local area network 24; and

c) it sends this data back to the archive service 16 via TCP/IP, by requesting the following URL:

http://www.example-url.net/d/?i=ZZZ&t=TTT&a=AAA

The archive service 16 receives the message which includes the parameters TTT, AAA, and ZZZ. The message also includes the IP address of the sender. This address is the customer's actual IP address, which in some cases is different from the HTTP proxy IP address. The archive service 16 records this information in the MDAS archive 17 and associates it with the transaction ID ZZZ.

The machine data collection and archiving process 1 of the present invention has been described with a particular application in fraud detection. However, it is foreseen that the techniques of the present invention have a wider application, as for marketing or computer support purposes, or other functions. While the process 1 has been described with reference to the internet 4 or world wide web, it is also conceivable that the process 1 could be employed on computer networks of less than global expanse, such as a large intranet, a national or regional network, or the like.

Therefore, it is to be understood that while certain forms of the present invention have been illustrated and described herein, the present invention is not intended to be limited to the specific forms, arrangement of parts, sequence of steps, or particular applications described and shown.

CLAIMS

What is claimed and desired to secure by Letters Patent is:

1. A process for collecting machine identifying information associated with a digital online access device used for substantially anonymously accessing a host computer system over a digital network, said host computer system generating an interaction record of an access therewith by said access device, and said process comprising the steps of:

   (a) capturing a machine identifying profile parameter upon said access device accessing said host computer system;

   (b) generating a unique interaction identification string upon said access device accessing said host computer system;

   (c) associating said interaction identification string with said profile parameter; and

   (d) associating said interaction identification string with said interaction record generated upon said access device accessing said host computer system.

2. A process as set forth in Claim 1 wherein said capturing step includes the step of:

   (a) capturing a digital address of said access device on said digital network.

3. A process as set forth in Claim 1 wherein said capturing step includes the step of:

   (a) capturing a configuration setting of said access device.

4. A process as set forth in Claim 1 and including the steps of:

(a) communicating a self-identification routine to said access device upon said access device accessing said host computer system;

(b) said access device executing said self-identification routine;

(c) said self-identification routine querying a configuration setting of said access device to derive said profile parameter; and

(d) said self-identification routine communicating said profile parameter to a remote location for association with said interaction identification string.

5. A process as set forth in Claim 1 and including the steps of:

(a) said host system operating a host web site including an interaction page generated by interaction page code processed by said access device upon accessing said host web site; and

(b) coding, within said interaction page code, a self-identification routine which causes said access device to communicate said profile parameter when said access device processes said interaction page code.

6. A process as set forth in Claim 3 and including the step of:

(a) coding said self-identification routine in such a manner that said profile parameter and said interaction identification string are communicated to a

third party web site at which said profile parameter and said interaction identification string are stored.

7. A process for identifying a customer computer involved in an online transaction between a customer using a customer browser operating on said customer computer and a merchant operating a merchant web site, said method comprising the steps of:

(a) capturing a customer computer profile parameter upon said customer computer accessing said merchant web site;

(b) generating a transaction identification string and associating said string with said parameter;

(c) storing said parameter and said string in a machine data archive;

(d) upon said customer completing a transaction through said merchant web site, storing said transaction identification string with a transaction record formed during said transaction to thereby associate said parameter with said transaction record through said string.

8. A process as set forth in Claim 7 wherein said capturing step includes the step of:

(a) capturing an IP address of said customer computer.

9. A process as set forth in Claim 7 wherein said capturing step includes the step of:

(a) capturing a configuration setting of said customer computer.

10. A process as set forth in Claim 7 and including the step of:

(a) communicating said profile parameter and said transaction identification string to a third party web site for storage.

11. A process as set forth in Claim 7 and including the step of:

(a) causing said customer computer to communicate said profile parameter and said transaction identification string to a third party web site for storage.

12. A process as set forth in Claim 7 and including the steps of:

(a) communicating a self-identification routine to said customer computer upon said customer computer accessing said merchant web site;

(b) said customer computer executing said self-identification routine;

(c) said self-identification routine querying a configuration setting of said customer computer to derive said profile parameter; and

(d) said self-identification routine communicating said profile parameter to a remote location for association with said interaction identification string.

13. A process as set forth in Claim 12 and including the step of:

(a) coding said self-identification routine in such a manner that said profile parameter and said interaction identification string are communicated to a third party web site at which said profile parameter and said interaction identification string are stored.

14. A process as set forth in Claim 12 wherein said querying step includes the steps of:

(a) querying said customer browser for a plurality of configuration settings;

(b) forming an attribute string from said plurality of configuration settings; and

(c) processing said attribute string to form said profile parameter of said customer computer.

15. A process as set forth in Claim 12 wherein said customer computer potentially accesses said merchant web site by way of a proxy, and said communicating step includes the steps of:

(a) communicating said profile parameter and said transaction identification string to said remote web site using a protocol which bypasses said proxy.

16. A process for identifying a customer computer involved in an online transaction through a merchant web site between a customer using a customer browser

operating on said customer computer and a merchant who operates said web site, said method comprising the steps of:

(a) coding a script request within a transaction form of said merchant web site;

(b) processing said script request by said customer browser upon accessing said merchant web site to thereby communicate to an archiver web site of a machine data archiving service an electronic request for a machine data collection script;

(c) said archiver web site returning said script to said customer browser along with a unique transaction identification string;

(d) said customer browser processing said script to thereby cause said script to query said customer computer for a profile parameter of said customer computer;

(e) said script causing said customer browser to communicate said profile parameter and said transaction identification string to said archiver web site;

(f) said archiver web site storing said profile parameter and said transaction identification string in a machine data profile;

(g) said script causing said customer browser to write said transaction identification string into said transaction form; and

(h) upon said customer adding customer identification information to said transaction form and electronically submitting said transaction form to said

merchant web site to thereby comprise a transaction record, said transaction identification string associating said transaction record with said machine data profile.

17. A process as set forth in Claim 16 and including the steps of:

(a) said script causing said computer browser to communicate said profile parameter and said transaction identification string along with a conventional hypertext transfer protocol (HTTP) header; and

(b) said archiver service additionally storing said HTTP header in association with said machine data profile.

18. A process as set forth in Claim 16 and including the step of:

(a) said script querying said customer browser for a configuration setting thereof.

19. A process as set forth in Claim 16 and including the steps of:

(a) said script querying said customer browser for a plurality of configuration settings;

(b) said script forming an attribute string from said plurality of configuration settings; and

(c) said script processing said attribute string to form said profile parameter of said customer computer.

20. A process as set forth in Claim 19 wherein said script processing step includes the step of:

(a) said script performing a hashing function on said attribute string to form said profile parameter.

21. A process as set forth in Claim 16 wherein said customer computer potentially accesses said merchant web site by way of a proxy, and including the step of:

(a) said script communicating said profile parameter and said transaction identification string to said archiver service web site using a protocol which bypasses said proxy.

22. A process as set forth in Claim 16 and including the step of:

(a) said script communicating said profile parameter to said archiver service web site using a protocol other than HTTP.

23. A process as set forth in Claim 16 wherein said customer computer includes a digital clock, and including the steps of:

(a) said script causing said customer browser to query said clock for a time value; and

(b) said script causing said customer browser to send said time value to said archiver service web site along with said profile parameter.

24. A process for identifying a customer computer involved in an online transaction through a merchant web site between a customer using a customer browser operating on said customer computer and a merchant who operates said web site, said method comprising the steps of:

(a) coding a script request within a transaction form of said merchant web site;

(b) processing said script request by said customer browser upon accessing said merchant web site to thereby communicate to an archiver web site of a machine data archiving service an electronic request for a machine data collection script;

(c) said archiver web site returning said script to said customer browser along with a unique transaction identification string;

(d) said customer browser processing said script to thereby cause said script to:

(1) query said customer browser for a plurality of configuration settings;

(2) form an attribute string from said plurality of configuration settings;

(3) perform a hashing function on said attribute string to form said profile parameter; and

(4) query an internal digital clock of said customer computer for a current time value;

(e) said script causing said customer browser to communicate said profile parameter, said time value, and said transaction identification string to said archiver web site along with a conventional HTTP header;

(f) said archiver web site storing said profile parameter, said time value, and said transaction identification string in a machine data profile;

(g) said script causing said customer browser to write said transaction identification string into said transaction form; and

(h) upon said customer adding customer identification information to said transaction form and electronically submitting said transaction form to said merchant web site to thereby comprise a transaction record, said transaction identification string associating said transaction record with said machine data profile.

25. A process as set forth in Claim 24 wherein said customer computer potentially accesses said merchant web site by way of a proxy, and including the steps of:

(a) said script querying said customer computer for a second profile parameter; and

(b) said script communicating said second profile parameter and said transaction identification string to said archiver service web site using a protocol which bypasses said proxy.

26. A process as set forth in Claim 25 and including the step of:

(a) said script communicating said second profile parameter to said archiver service web site using a protocol other than HTTP.

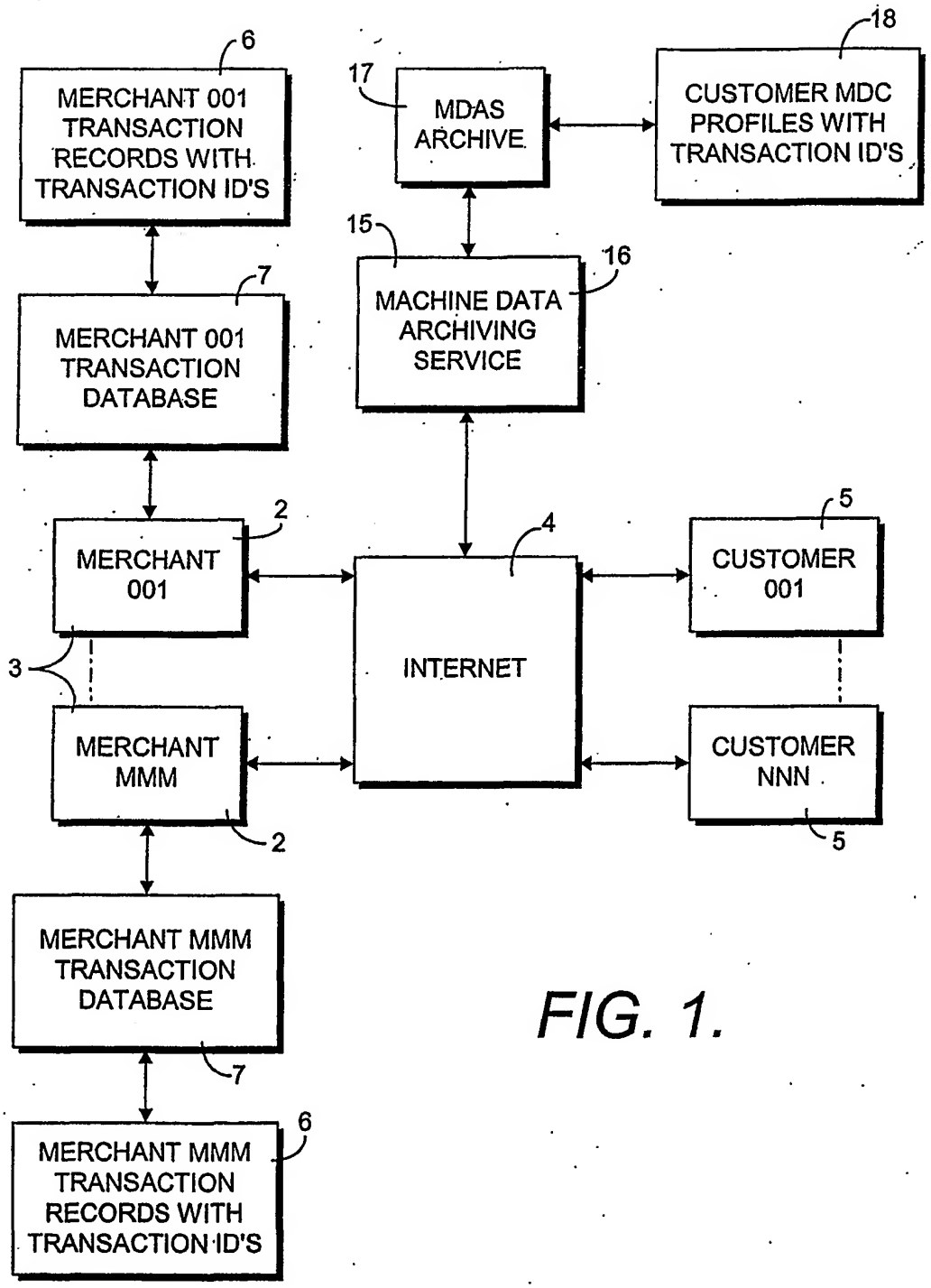

6
MERCHANT 001 TRANSACTION RECORDS WITH TRANSACTION ID'S
17
MDAS ARCHIVE
18
CUSTOMER MDC PROFILES WITH TRANSACTION ID'S
15
16
MACHINE DATA ARCHIVING SERVICE
7
MERCHANT 001 TRANSACTION DATABASE
2
MERCHANT 001
4
INTERNET
5
CUSTOMER 001
3
MERCHANT MMM
2
CUSTOMER NNN
5
MERCHANT MMM TRANSACTION DATABASE
7
6
MERCHANT MMM TRANSACTION RECORDS WITH TRANSACTION ID'S

FIG. 1.

FIG. 2.

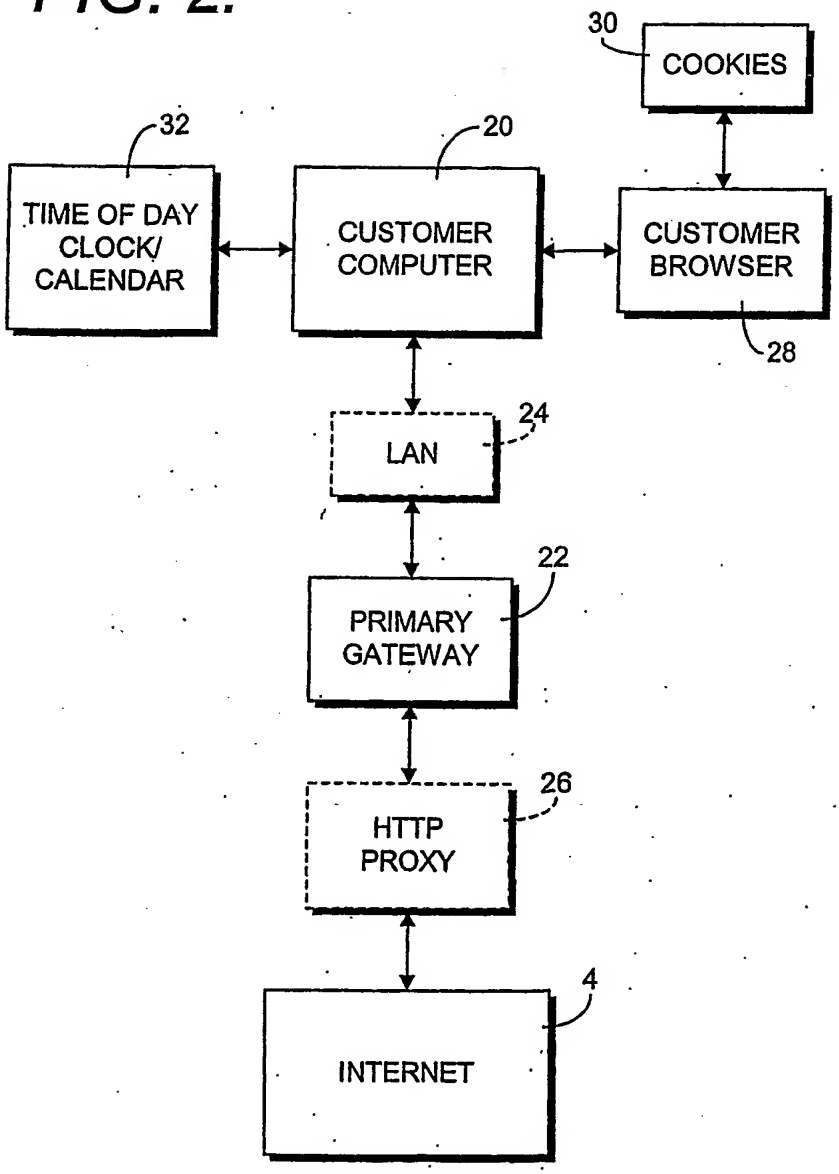

30
COOKIES
32
20
TIME OF DAY CLOCK/ CALENDAR
CUSTOMER COMPUTER
CUSTOMER BROWSER
28
24
LAN
22
PRIMARY GATEWAY
26
HTTP PROXY
4
INTERNET

FIG. 3.

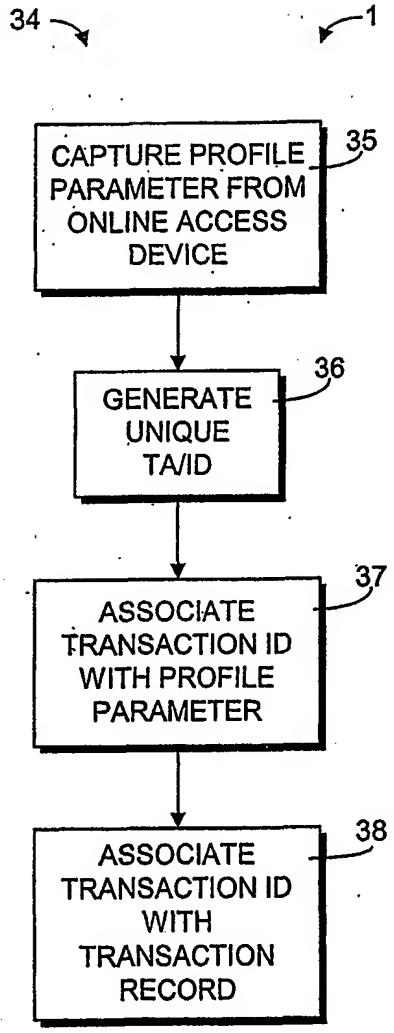
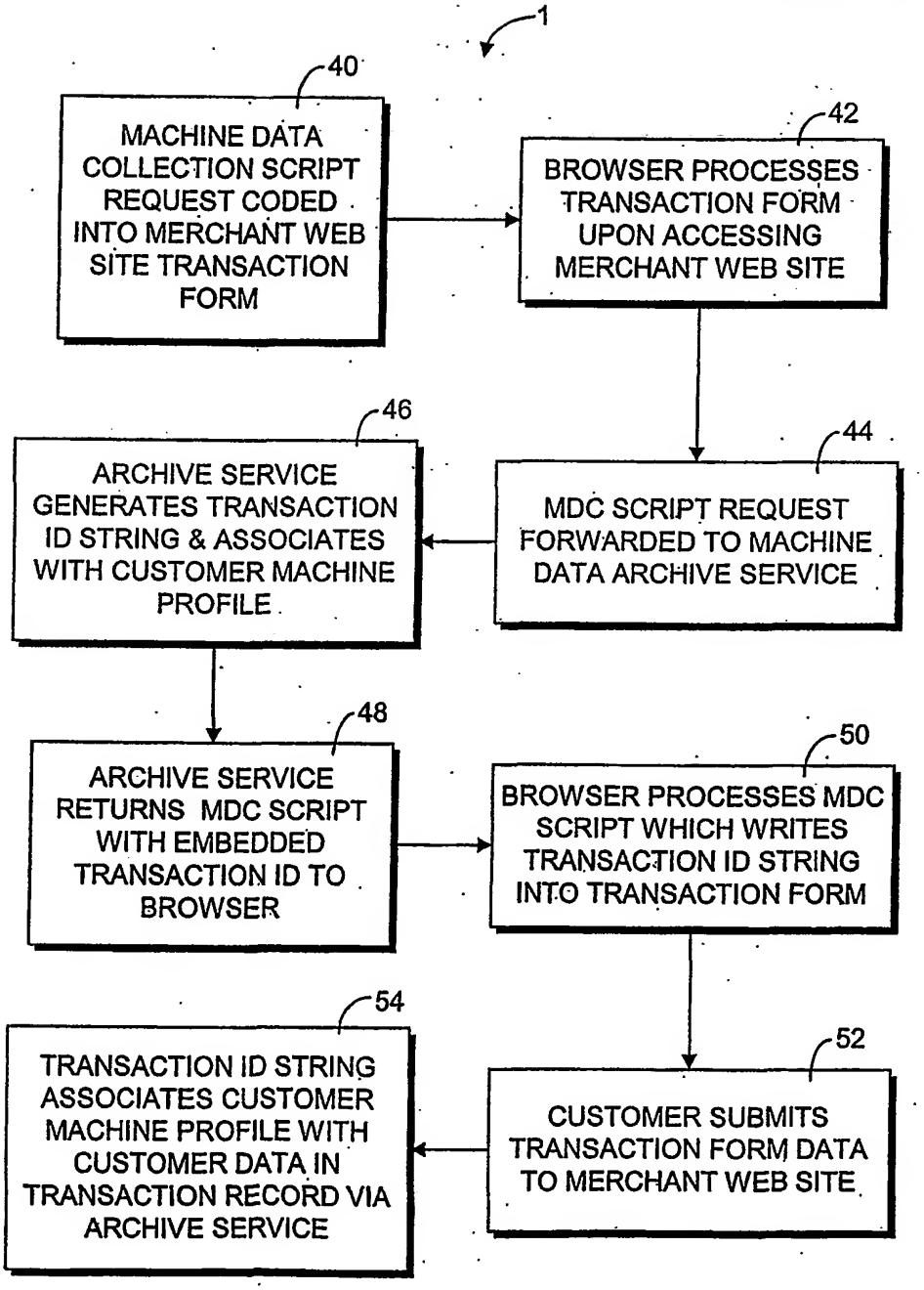

34
1
CAPTURE PROFILE PARAMETER FROM ONLINE ACCESS DEVICE
35
GENERATE UNIQUE TA/ID
36
ASSOCIATE TRANSACTION ID WITH PROFILE PARAMETER
37
ASSOCIATE TRANSACTION ID WITH TRANSACTION RECORD
38

FIG. 4.


1
40
MACHINE DATA COLLECTION SCRIPT REQUEST CODED INTO MERCHANT WEB SITE TRANSACTION FORM
42
BROWSER PROCESSES TRANSACTION FORM UPON ACCESSING MERCHANT WEB SITE
44
MDC SCRIPT REQUEST FORWARDED TO MACHINE DATA ARCHIVE SERVICE
46
ARCHIVE SERVICE GENERATES TRANSACTION ID STRING & ASSOCIATES WITH CUSTOMER MACHINE PROFILE
48
ARCHIVE SERVICE RETURNS MDC SCRIPT WITH EMBEDDED TRANSACTION ID TO BROWSER
50
BROWSER PROCESSES MDC SCRIPT WHICH WRITES TRANSACTION ID STRING INTO TRANSACTION FORM
52
CUSTOMER SUBMITS TRANSACTION FORM DATA TO MERCHANT WEB SITE
54
TRANSACTION ID STRING ASSOCIATES CUSTOMER MACHINE PROFILE WITH CUSTOMER DATA IN TRANSACTION RECORD VIA ARCHIVE SERVICE

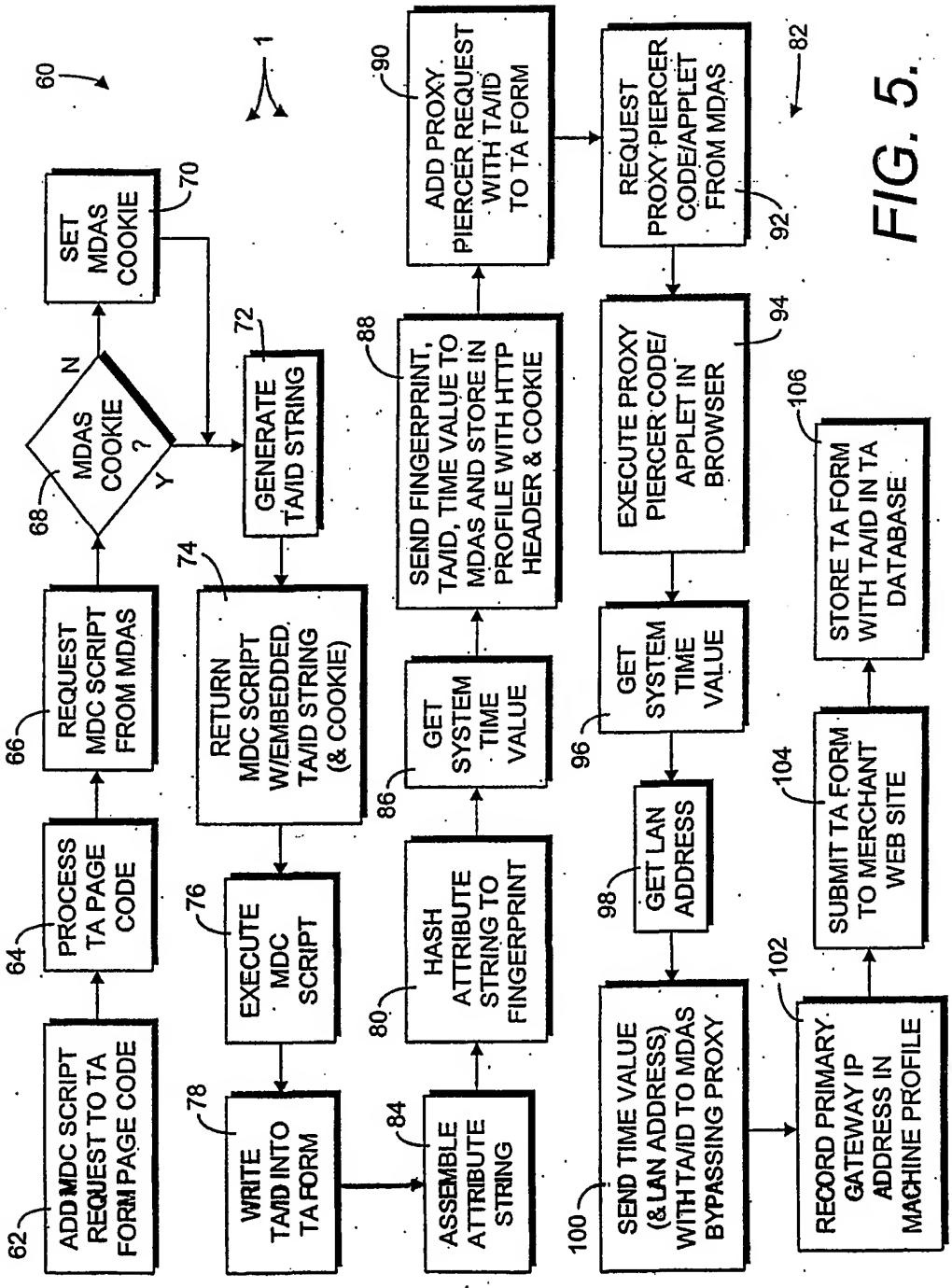

60
1
ADD MDC SCRIPT REQUEST TO TA FORM PAGE CODE
62
PROCESS TA PAGE CODE
64
REQUEST MDC SCRIPT FROM MDAS
66
MDAS COOKIE ?
68
N
Y
SET MDAS COOKIE
70
GENERATE TA/ID STRING
72
RETURN MDC SCRIPT W/EMBEDDED TA/ID STRING (& COOKIE)
74
EXECUTE MDC SCRIPT
76
WRITE TA/ID INTO TA FORM
78
HASH ATTRIBUTE STRING TO FINGERPRINT
80
ASSEMBLE ATTRIBUTE STRING
84
GET SYSTEM TIME VALUE
86
SEND FINGERPRINT, TA/ID, TIME VALUE TO MDAS AND STORE IN PROFILE WITH HTTP HEADER & COOKIE
88
ADD PROXY PIERCER REQUEST WITH TA/ID TO TA FORM
90
REQUEST PROXY PIERCER CODE/APPLET FROM MDAS
92
82
EXECUTE PROXY PIERCER CODE/ APPLET IN BROWSER
94
GET SYSTEM TIME VALUE
96
GET LAN ADDRESS
98
SEND TIME VALUE (& LAN ADDRESS) WITH TA/ID TO MDAS BYPASSING PROXY
100
RECORD PRIMARY GATEWAY IP ADDRESS IN MACHINE PROFILE
102
SUBMIT TA FORM TO MERCHANT WEB SITE
104
STORE TA FORM WITH TA/ID IN TA DATABASE
106

FIG. 5.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/18076

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : G06F 17/60
US CL : 705/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 705/10

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EAST search terms: digital, anonymous, server, network, web, record, profile, interaction, transaction, identification, string, proxy, third party

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5,848,412 A (ROWLAND et al) 08 December 1998, all.	1-26
A	US 5,991,758 A (ELLARD) 23 November 1999, all.	1-26
A.P	US 6,117,011 A (LVOV) 12 September 2000, all.	1-26

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

06 AUGUST 2001

Date of mailing of the international search report

04 SEP 2001

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

Tariq Hafiz

Telephone No. (703) 305-3900